# Distributionally robust multistage inventory models with moment constraints

**Linwei Xin**
School of Industrial & Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332
**David A. Goldberg**
School of Industrial & Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332
**Alexander Shapiro***
School of Industrial & Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332

In this paper, we consider a minimax approach to managing an inventory under distributional uncertainty. In particular, we study the associated multistage distributionally robust optimization problem, when only the mean, variance, and distribution support are known for the demand at each stage. It is known that if the policy maker is allowed to recompute her policy choice after each stage (i.e. dynamic formulation), thus taking prior realizations of demand into consideration when performing the relevant minimax calculations at later stages, a basestock policy is optimal. In contrast, if the policy maker is not allowed to recompute her policy after each stage (i.e. static formulation), far less is known. If these two formulations have a common optimal policy, i.e. the policy maker would be content with the given policy whether or not she has the power to recompute after each stage, we say that the policy is *time consistent*, and the problem is *weakly time consistent*. If every optimal policy for the static formulation is time consistent, we say that the problem is *strongly time consistent*. In this paper, we give sufficient conditions for weak and strong time consistency. We also provide several examples demonstrating that the problem is not time consistent in general. Furthermore, these examples show that the question of time consistency can be quite subtle in this setting. In particular, we show that: (i) the problem can fail to be weakly time consistent, (ii) the problem can be weakly but not strongly time consistent, and (iii) the problem can be strongly time consistent even if the two formulations have different optimal values. Interestingly, this stands in contrast to the analogous setting in which only the mean and support of the demand distribution is known at each stage, for which it is known that such time inconsistency cannot occur Shapiro (2012).

*Key words*: inventory model, distributional robustness, moment constraints, multistage program, dynamic programming, basestock policy, time consistency, backorder and holding costs

## 1. Introduction

The news vendor (news boy) problem, used to analyze the trade-offs associated with stocking an inventory, has its origin in a seminal paper by Edgeworth (1888). In its classical formulation, the

* Research of this author was partly supported by the NSF award CMMI 1232623.

problem is stated as a minimization of the expected value of the relevant ordering, backorder, and holding costs. Such a formulation requires a complete specification of the probability distribution of the underlying demand process. However, in applications knowledge of the exact distribution of the demand process is rarely available. This motivates the study of minimax type (i.e. distributionally robust) formulations, where minimization is performed with respect to a worst-case distribution from some family of potential distributions (cf. Dupacová (1987, 2001), Prékopa (1995), Žáčková (1966)). In a pioneering paper Scarf (1958) gave an elegant solution for the minimax news vendor problem when only the first and second moments of the demand distribution are known. His work has led to considerable follow-up work, and we refer the reader to Gallego and Moon (1993) and Gallego, Katircioglu and Ramachandran (2007), and references therein, for related results. For a more general overview of risk analysis for news vendor and inventory models we can refer, e.g., to Ahmed, Cakmak and Shapiro (2007) and Choi, Ruszczyński and Zhao (2011).

In practice an inventory must often be managed over some time horizon, and the classical news vendor problem was naturally extended to the multistage setting, for which there is also a considerable literature (see Zipkin (2000) and references therein). Recently, robust variants of such multistage problems have begun to receive attention in the literature. It has been observed that such multistage robust optimization problems can exhibit a subtle phenomenon known as time inconsistency (e.g., Artzner et al. (2007), Carpentier et al. (2012), Huang et al. (2011), Riedel (2004), Ruszczyński (2010), Shapiro (2009)).

Various formal definitions of time consistency have been proposed in the literature. We will give a formal definition, which is naturally suited to our framework, in Section 4. The intuition is as follows. A multistage distributionally robust optimization problem can be viewed in two ways. In one formulation, the policy maker is allowed to recompute her policy choice after each stage (i.e. dynamic formulation), thus taking prior realizations of demand into consideration when performing the relevant minimax calculations at later stages, in which case it is known that a basestock policy is optimal. In the second formulation, the policy maker is not allowed to recompute her policy after each stage (i.e. static formulation), in which case far less is known. If these two formulations have a common optimal policy, i.e. the policy maker would be content with the given policy whether or not she has the power to recompute after each stage, we say that the policy is *time consistent*, and the problem is *weakly time consistent*. If every optimal policy for the static formulation is time consistent, we say that the problem is *strongly time consistent*. Such a property is desirable from a policy perspective, as it ensures that previously agreed upon policy decisions still make sense when the policy is actually implemented, possibly at a later time. This approach is in line with the original work of Bellman (1957) on dynamical programming. In the context of distributionally robust formulations for the news vendor problem, it was recently shown that if all one knows about

the demand at each stage is the support and first moment, then time inconsistency cannot occur, in an appropriate sense (cf., Shapiro 2012). However, beyond these results, very little is known.

In this paper, we extend the work of Scarf (1958) by considering the multistage inventory (news vendor) problem when the support and first two moments are known for the demand at each stage. We give sufficient conditions for weak and strong time consistency. We also provide several examples demonstrating that the problem is not time consistent in general. Furthermore, these examples show that the question of time consistency can be quite subtle in this setting. In particular, we show that: (i) the problem can fail to be weakly time consistent, (ii) the problem can be weakly but not strongly time consistent, and (iii) the problem can be strongly time consistent even if the two formulations have different optimal values. Interestingly, this stands in contrast to the analogous setting in which only the mean and support of the demand distribution is known at each stage, for which it is known that such time inconsistency cannot occur Shapiro (2012).

The structure of the rest of the paper is as follows. In Section 2, we review the setup and known results for the single stage setting, including the results of Scarf (1958), and provide new results, which extend Scarf's analysis of the single stage distributionally robust news vendor problem to a new family of objective functions. In Section 3, we formally introduce the multistage distributionally robust news vendor problem, and review the static and dynamic formulations. In Section 4, which comprises the main results of the paper, we explore the notion of time consistency for the multistage distributionally robust news vendor problem. In Subsection 4.1 we give sufficient conditions for weak and strong time consistency. In Subsection 4.2 we give several examples, which demonstrate that the problem of interest is not time consistent in general, and further show that the question of time consistency can be quite subtle. We provide concluding remarks and directions for future research in Section 5. Also, we include a technical appendix in Section 6.

## 2. Single stage formulation

In this section we review both the classical and distributionally robust single stage formulation, including some relevant results of Scarf (1958). We also extend Scarf's approach to provide a new explicit solution for a particular class of objective functions, which we will later use in our study of time consistency.

### 2.1. Classical formulation

Consider the following classical formulation of the news vendor problem:

$$\min_{x \geq 0} \mathbb{E}[\Psi(x, D)], \tag{1}$$

where

$$\Psi(x, d) := cx + b[d - x]_+ + h[x - d]_+, \tag{2}$$

and $c, b, h$ are the ordering, backorder penalty, and holding costs, per unit, respectively. Unless stated otherwise we assume that $b > c > 0$ and $h \geq 0$. The expectation is taken with respect to the probability distribution of the demand $D$, which is modeled as a random variable (r.v.) having nonnegative support. It is well known that this problem has the closed form solution $\bar{x} = F^{-1}\left(\frac{b-c}{b+h}\right)$, where $F(\cdot)$ is the cumulative distribution function (cdf) of the demand $D$, and $F^{-1}$ is its inverse. Of course, it is assumed here that the probability distribution, i.e. the cdf $F$, is completely specified.

### 2.2. Distributionally robust formulation

Suppose now that the probability distribution of the demand $D$ is not fully specified, but instead assumed to be a member of a family of distributions $\mathfrak{M}$. Then we instead consider the following distributionally robust formulation:

$$\min_{x \geq 0} \psi(x), \tag{3}$$

where

$$\psi(x) := \sup_{Q \in \mathfrak{M}} \mathbb{E}_Q[\Psi(x, D)], \tag{4}$$

and the notation $\mathbb{E}_Q$ emphasizes that the expectation is taken with respect to the distribution $Q$ of the demand $D$.

We now introduce some additional notations to describe certain families of distributions. For a probability measure (distribution) $Q$, we let $\text{supp}(Q)$ denote the support of the measure, i.e. the smallest closed set $A \subset \mathbb{R}$ such that $Q(A) = 1$. With a slight abuse of notation, for a r.v. $Z$, we also let $\text{supp}(Z)$ denote the support of the associated probability measure. For a given closed (and possibly unbounded) interval $\mathcal{I} \subset \mathbb{R}$, we let $\mathfrak{P}(\mathcal{I})$ denote the set of probability distributions $Q$ such that $\text{supp}(Q) \subset \mathcal{I}$. Although we will be primarily interested in the setting that $\mathcal{I} \subseteq \mathbb{R}_+$ (i.e. demand is nonnegative), it will sometimes be convenient for us to consider more general families of demand distributions. By $\delta_a$ we denote the probability measure of mass one at $a \in \mathbb{R}$.

In this paper, we will study families of distributions satisfying moment constraints of the form

$$\mathfrak{M} := \left\{ Q \in \mathfrak{P}(\mathcal{I}) : \mathbb{E}_Q[D] = \mu, \mathbb{E}_Q[D^2] = \mu^2 + \sigma^2 \right\}. \tag{5}$$

Unless stated otherwise, it will be assumed that $\mathfrak{M}$ is indeed of the form (5), and is nonempty. We let $\alpha$ denote the left-endpoint of $\mathcal{I}$ (or $-\infty$ if $\mathcal{I}$ is unbounded from below), and let $\beta$ denote the right-endpoint of $\mathcal{I}$ (or $\infty$ if $\mathcal{I}$ is unbounded from above); i.e., $\mathcal{I} = [\alpha, \beta]$. It may be easily verified that, in case of bounded interval $[\alpha, \beta]$, the set $\mathfrak{M}$ is nonempty iff the following conditions hold

$$\mu \in [\alpha, \beta] \text{ and } \sigma^2 \leq (\beta - \mu)(\mu - \alpha), \tag{6}$$

which will be assumed throughout. Furthermore, one can also identify conditions under which $\mathfrak{M}$ is a singleton.

**Observation 1** *If $-\infty < \alpha < \beta < \infty$, $\mu \in [\alpha, \beta]$, and $\sigma^2 = (\beta - \mu)(\mu - \alpha)$, then $\mathfrak{M}$ consists of the single probability measure which assigns to the point $\alpha$ probability $p = \frac{\beta - \mu}{\beta - \alpha}$, and to the point $\beta$ probability $1 - p = \frac{\mu - \alpha}{\beta - \alpha}$.*

Our definitions imply that for all $x \in \mathbb{R}$, $\psi(x)$ equals the optimal value of the following maximization problem:

$$\begin{array}{ll} \max\limits_{Q \in \mathfrak{P}(\mathcal{I})} & \int \Psi(x, \tau) dQ(\tau) \\ \text{s.t.} & \int \tau dQ(\tau) = \mu, \ \int \tau^2 dQ(\tau) = \mu^2 + \sigma^2. \end{array} \tag{7}$$

Problem (7) is a classical problem of moments (see, e.g., Landau 1987). By the Richter-Rogosinski Theorem (e.g., Shapiro, Dentcheva and Ruszczyński 2009, Proposition 6.40) we have the following.

**Observation 2** *If Problem (7) possesses an optimal solution, then it has an optimal solution with support on at most three points.*

We note that the distributionally robust single stage news vendor problem considered by Scarf (1958) is exactly Problem (3) when the objective function $\Psi(x, d)$ is of the form (2), and $\mathcal{I} = \mathbb{R}_+$. As it will be useful for later proofs, we briefly review Scarf's explicit solution. We actually state a slight generalization of the results of Scarf, and for completeness we include a proof in Section 6.

**Theorem 1** *Suppose that $b > c$, $c + h > 0$, $\mu > 0$, $\sigma > 0$, and $\mathcal{I} = \mathbb{R}_+$. Let $\kappa := \frac{b-h-2c}{b+h}$. Then for each $x \in \mathbb{R}$,*

$$\psi(x) = \begin{cases} c\mu + \frac{b+h}{2}\left((x-\mu)^2 + \sigma^2\right)^{\frac{1}{2}} - \frac{b-h-2c}{2}(x-\mu), & \text{if } x \geq \frac{\mu^2+\sigma^2}{2\mu}, \\ \frac{(h+c)\sigma^2 - (b-c)\mu^2}{\mu^2+\sigma^2} x + b\mu, & \text{if } x \in [0, \frac{\mu^2+\sigma^2}{2\mu}), \\ b\mu - (b-c)x, & \text{otherwise}. \end{cases} \tag{8}$$

*As a consequence, a complete solution to the problem $\min_{x \in \mathbb{R}} \psi(x)$ is as follows.*
(i) *If $\frac{\sigma^2}{\mu^2} > \frac{b-c}{h+c}$, then the unique optimal solution is $x = 0$, and the optimal value is $\mu b$.*
(ii) *If $\frac{\sigma^2}{\mu^2} < \frac{b-c}{h+c}$, then the unique optimal solution is $x = \mu + \kappa\sigma(1-\kappa^2)^{-\frac{1}{2}}$, and the optimal value is $c\mu + \left((h+c)(b-c)\right)^{\frac{1}{2}}\sigma$.*
(iii) *If $\frac{\sigma^2}{\mu^2} = \frac{b-c}{h+c}$, then all $x \in [0, \mu + \kappa\sigma(1-\kappa^2)^{-\frac{1}{2}}]$ are optimal, and the optimal value is $\mu b$.*
*Furthermore, we note that in all cases, for all $x \in \mathbb{R}$, $\arg\max_{Q \in \mathfrak{M}} \mathbb{E}_Q[\Psi(x, D)]$ is non-empty. Also, the optimal solution set and value of the problem $\min_{x \in \mathbb{R}} \psi(x)$ is identical to that of Problem (3), i.e. optimizing over $x \in \mathbb{R}$, as opposed to $x \in \mathbb{R}_+$, makes no difference.*

For use in later proofs, it will also be useful to demonstrate a particular variant of Theorem 1. Suppose that in Problem (3), we are not forced to select a deterministic constant $x$, but can instead select any distribution $D_1$ for $x$. Specifically, let us consider the following minimax problem:

$$\min_{Q_1 \in \mathfrak{P}(\mathcal{I})} \phi(Q_1), \tag{9}$$

where

$$\phi(Q_1) := \max_{Q_2\in\mathfrak{M}} \mathbb{E}_{Q_1\times Q_2}\big[\Psi(D_1,D_2)\big],$$

and the notation $\mathbb{E}_{Q_1\times Q_2}$ indicates that for any choices for the marginal distributions $Q_1, Q_2$ of $D_1$ and $D_2$, the expectation is taken w.r.t. the associated product measure, under which $D_1$ and $D_2$ are independent. In this case, we have the following result.

**Proposition 2.1** *Suppose that $b > c$, $c + h > 0$, $\mu > 0$, $\sigma > 0$, $\frac{\sigma^2}{\mu^2} > \frac{b-c}{h+c}$, and $\mathcal{I} = \mathbb{R}$. Then Problem* (9) *has the unique optimal solution $\bar{Q}_1 = \delta_0$.*

*Proof* Since problem (9) has no constraints, it suffices to optimize over measures of mass one at a point $x \in \mathbb{R}$. Of course, for $Q_1 = \delta_x$ we have that $\phi(Q_1) = \psi(x)$. Therefore $\bar{Q}_1 = \delta_0$ is an optimal solution of (9) by Theorem 1. For a proof of uniqueness we refer to the Appendix (Section 6) . □

We also note that $\psi$ inherits the property of convexity from $\Psi$.

**Observation 3** *If $\Psi(\cdot, d)$ is a convex function for every fixed $d \in \mathcal{I}$, then $\psi$ is also a convex function, and Problem* (3) *is a convex program.*

As several of our later proofs will be based on duality theory, we now briefly review duality for Problem (7).

### 2.3. Duality for Problem (7)

The dual of Problem (7) can be constructed as follows (cf., Isii 1963). Consider the Lagrangian

$$L(Q,\lambda) := \int \Big[\Psi(x,\tau) - \sum_{i=0}^{2} \lambda_i \tau^i\Big] dQ(\tau) + \lambda_0 + \lambda_1\mu + \lambda_2(\mu^2+\sigma^2).$$

By maximizing $L(Q,\lambda)$ with respect to $Q \in \mathfrak{P}(\mathcal{I})$, and then minimizing with respect to $\lambda$, we obtain the following Lagrangian dual for Problem (7):

$$\begin{array}{ll} \min\limits_{\lambda\in\mathbb{R}^3} & \lambda_0 + \lambda_1\mu + \lambda_2(\mu^2+\sigma^2) \\ \text{s.t.} & \lambda_0 + \lambda_1\tau + \lambda_2\tau^2 \ge \Psi(x,\tau), \ \ \tau \in \mathcal{I}. \end{array} \tag{10}$$

We denote by $\mathrm{val}(P)$ and $\mathrm{val}(D)$ the respective optimal values of the primal Problem (7) and its dual Problem (10). By convention, if Problem (7) is infeasible, we set $\mathrm{val}(P) = -\infty$, and if Problem (10) is infeasible, we set $\mathrm{val}(D) = \infty$. We denote by $\mathrm{Sol}_P(x)$ the set of optimal solutions of the primal problem, and by $\mathrm{Sol}_D(x)$ the set of optimal solutions of the dual problem, and note that these sets may be empty, even when both programs are feasible, e.g. if the respective optimal values are approached but not attained.

Note that whenever $P$ is feasible (which we assume throughout), $\mathrm{val}(P) = \psi(x)$, and $\mathrm{val}(D) \ge \mathrm{val}(P)$. We now give sufficient conditions for there to be no duality gap, i.e. $\mathrm{val}(P) = \mathrm{val}(D)$, as

well as conditions for Problems (7) and (10) to have optimal solutions. By specifying known general results for duality of such programs, e.g., (Shapiro 2001, Corollary 3.1, Propositions 3.4 and 3.5) and (Bonnans and Shapiro 2000, Theorem 5.97), to the considered setting, we have the following.

**Proposition 2.2** *If $\bar{Q}$ is a probability measure which is feasible for the primal Problem* (7), $\bar{\lambda} = (\bar{\lambda}_0, \bar{\lambda}_1, \bar{\lambda}_2)$ *is a vector which is feasible for the dual Problem* (10), *and*

$$\operatorname{supp}(\bar{Q}) \subseteq \left\{\tau \in \mathcal{I} : \Psi(x,\tau) = \bar{\lambda}_0 + \bar{\lambda}_1 \tau + \bar{\lambda}_2 \tau^2\right\}, \tag{11}$$

*then $\bar{Q}$ is an optimal primal solution, $\bar{\lambda}$ is an optimal dual solution, and* $\operatorname{val}(P) = \operatorname{val}(D)$. *Conversely, if* $\operatorname{val}(P) = \operatorname{val}(D)$, *and $\bar{Q}$ and $\bar{\lambda}$ are optimal solutions of the respective primal and dual problems, then Condition* (11) *holds.*

## 2.4. Explicit solution of Problem (7) for a class of convex, continuous, piecewise affine functions

Scarf (1958) gave an explicit solution for Problems (7) and (10) when $\mathcal{I} = \mathbb{R}_+$, and $\Psi(x,\cdot)$ is a convex, continuous piecewise affine function with exactly two pieces, by explicitly constructing a feasible primal - dual solution pair satisfying the conditions of Proposition 2.2 (details of this construction can be found in Section 6). Here, we generalize Scarf's approach to a class of convex, continuous, piecewise affine (CCPA) functions with three pieces, as we will need the solution to such problems for our later studies of time consistency. We note that Scarf's approach can also be extended to the family of general CCPA functions with a general number of pieces, although we do not pursue that here. In particular, we establish the following result, and defer the proof to the Appendix (Section 6).

**Theorem 2** *Suppose that for some $x \in \mathbb{R}$, there exist $c_1, c_2 > 0$ s.t. $c_1 < c_2$, and* $\Psi(x,d) = \max\{-d + c_1, 0, d - c_2\}$ *for all $d \in \mathbb{R}$. Let $\eta := \frac{1}{2}(c_1 + c_2)$, and $\mathfrak{f}(z) := \left((z-\mu)^2 + \sigma^2\right)^{\frac{1}{2}}$ for all $z \in \mathbb{R}$. Further suppose that $\sigma > 0$, $\mathcal{I} = \mathbb{R}_+$,*

$$\frac{1}{4}(2\mu - 3c_1 + c_2)(3c_2 - c_1 - 2\mu) \le \sigma^2,$$

*and $\eta - \mathfrak{f}(\eta) \ge 0$. Then the unique optimal solution to the primal Problem* (7) *is the probability measure $Q$ having support at two points $\eta - \mathfrak{f}(\eta)$ and $\eta + \mathfrak{f}(\eta)$, with*

$$Q\big(\eta - \mathfrak{f}(\eta)\big) = \sigma^2 \left(\sigma^2 + \big(\eta - \mathfrak{f}(\eta) - \mu\big)^2\right)^{-1}, \ Q\big(\eta + \mathfrak{f}(\eta)\big) = 1 - Q\big(\eta - \mathfrak{f}(\eta)\big). \tag{12}$$

*Also, the unique optimal solution to the dual Problem* (10) *is*

$$\lambda_0 = \tfrac{1}{2}\big(\eta^2 + (\eta-\mu)^2 + \sigma^2\big)\mathfrak{f}^{-1}(\eta) + \frac{c_1 - c_2}{2}, \quad \lambda_1 = -\eta \mathfrak{f}^{-1}(\eta), \quad \lambda_2 = \tfrac{1}{2}\mathfrak{f}^{-1}(\eta). \tag{13}$$

## 3. Multistage formulation

In this section, we study a multistage extension of the distributionally robust news vendor problem discussed in Section 2.2.

### 3.1. Classical formulation

We begin by reviewing the classical (i.e. non-robust) multistage news vendor problem, and start by introducing some additional notations. For a vector $z \in \mathbb{R}^n$, and $k \in \{1, ..., n\}$, let $z_{[k]}$ denote the vector consisting of the first $k$ components of $z$, i.e. $z_{[k]} := (z_1, \dots, z_k)$, and $z_{[0]} := \emptyset$.

We suppose that there is a finite time horizon $T$, and a (random) vector of demands $D = (D_1, \dots, D_T)$, such that $\{D_t,\ t = 1, \dots, T\}$ are *mutually independent.* We now define the family of admissible policies $\Pi$ by introducing two families of functions, $y = \{y_t,\ t = 1, \dots, T\}$ and $x = \{x_t,\ t = 1, \dots, T\}$. Conceptually, $y_t$ will correspond to the inventory level at the start of stage $t$, and $x_t(y_t)$ will correspond to the inventory level after having ordered in stage $t$, but before the demand in that stage is realized.

Note that for $t \geq 2$, $y_t$ is a function of $D_{[t-1]}$, and we restrict ourselves to policies such that the amount ordered in stage $t$ depends on realizations of past demand only through the inventory level at the start of stage $t$, i.e. the amount ordered in stage $t$ is a function of $y_t$ rather than the whole history $D_{[t-1]}$ of the demand process (justification for this assumption will be given below). Such policies are nonanticipative, i.e. decisions do not depend on realizations of future demand. We assume that $y_1$, the initial inventory level, is a given constant. We also require that one can only order a nonnegative amount of inventory at each stage. It follows that the set of admissible policies $\Pi$ should consist of those vectors of one-dimensional (measurable) functions $\pi = (x_1(y_1), x_2(y_2), \dots, x_T(y_T))$ such that $x_t : \mathbb{R} \to \mathbb{R}$ satisfies $x_t(y) \geq y$, for all $y \in \mathbb{R}$ and $t = 1, ..., T$. To be consistent with the inventory dynamics, we also require that

$$y_{t+1} = x_t(y_t) - D_t,\ t = 1, ..., T-1. \tag{14}$$

It follows that any given choice of $\pi \in \Pi$, along with the given $y_1$, completely determines the associated functions $y_1, \dots, y_T$. Combining the above, we can write the classical multistage inventory (news vendor) problem as follows:

$$\min_{\pi \in \Pi} \mathbb{E} \left\{ \sum_{t=1}^{T} \rho^{t-1} \big[ c_t \big( x_t(y_t) - y_t \big) + \Psi_t \big( x_t(y_t), D_t \big) \big] \right\}. \tag{15}$$

Here $\rho \in (0, 1]$ is a discount factor, $c_t, b_t, h_t$ are the ordering, backorder penalty and holding costs per unit in stage $t$, respectively, and

$$\Psi_t(x_t, d_t) := b_t[d_t - x_t]_+ + h_t[x_t - d_t]_+. \tag{16}$$

Unless stated otherwise we assume that $b_t > c_t > 0$ and $h_t \geq 0$ for all $t = 1, ..., T$.

With Problem (15) are associated the following dynamic programming equations (e.g., Zipkin (2000))

$$V_t(y) = \min_{x \geq y} \left\{ c_t(x-y) + \mathbb{E}[\Psi_t(x, D_t) + \rho V_{t+1}(x - D_t)] \right\}, \tag{17}$$

$t = 1, ..., T$, with $V_{T+1}(\cdot) \equiv 0$. The dynamic Equations (17) naturally define a set of optimal policies through the relation

$$x_t(y) \in \mathfrak{X}_t(y) := \arg\min_{x \geq y} \left\{ c_t(x-y) + \mathbb{E}[\Psi_t(x, D_t) + \rho V_{t+1}(x - D_t)] \right\},\ t = 1, ..., T, \tag{18}$$

with associated optimal value given by $V_1(y_1)$. Note that $x_t(y)$ are (measurable) functions of $y$, $t = 1, \ldots, T$.

It is well-known that the two formulations (15) and (17) are equivalent in the following sense. The optimal values of these problems are equal to each other, i.e. both have value $V_1(y_1)$, and the respective sets of optimal policies are the same for both formulations. More precisely, for any optimal policy $\bar{\pi} = (\bar{x}_1, \ldots, \bar{x}_T)$ for Problem (15), and any $t \in [1, T]$, there exists a set $A \in \mathbb{R}$ s.t. $P(y_t \in A) = 1$, and $x_t(y) \in \mathfrak{X}_t(y)$ for all $y \in A$. As we shall see, this equivalence does not necessarily hold for distributionally robust multistage inventory problems with moment constraints.

Note that it follows from (18) that the optimal policies for Problem (15) indeed have the property that the amount ordered in stage $t$ is a function of $y_t$, and thus there is no loss of generality in considering only policies of that form. Of course, the assumption of mutual (stage-wise) independence is essential for this conclusion.

### 3.2. Distributionally robust formulations

Suppose now that the distribution of the demand process is not known, and we only have at our disposal information about the support and first and second moments. In this case, it is natural to use the framework previously developed for the single stage problem (see Section 2) to handle the distributional uncertainty at each stage. However, in the multistage setting, there is a nontrivial question of how to model the associated uncertainty in the joint distribution of demand. The approach we take here is similar to that taken in previous studies of time consistency of the inventory model (cf., Shapiro 2012). We will consider two formulations, one intuitively corresponding to the modeling choices of a policy maker who does not recompute her policy choices after each stage (referred to as the *static formulation*), and one corresponding to a policy-maker who does (referred to as the *dynamic formulation*). We suppose that we have been given a sequence of closed (possibly unbounded) intervals $\mathcal{I}_t = [\alpha_t, \beta_t] \subset \mathbb{R}$, $t = 1, \ldots, T$, and sequences of the corresponding means $\mu_t$, $t = 1, \ldots, T$, and variances $\sigma_t^2$, $t = 1, \ldots, T$.

**3.2.1. Static formulation** We first consider the following formulation, referred to as *static*, in which the policy maker does not recompute her policy choices after each stage. Let us define

$$\mathfrak{M}_t := \left\{Q_t \in \mathfrak{P}(\mathcal{I}_t) : \mathbb{E}_{Q_t}[D_t] = \mu_t,\ \mathbb{E}_{Q_t}[D_t^2] = \mu_t^2 + \sigma_t^2\right\},\ t = 1, ..., T, \tag{19}$$

$$\mathfrak{M} := \{Q = Q_1 \times \cdots \times Q_T : Q_t \in \mathfrak{M}_t,\ t = 1, ..., T\}, \tag{20}$$

i.e. $\mathfrak{M}$ is the set of all product measures whose marginals belong to the associated sets $\mathfrak{M}_t$, $t = 1, \ldots, T$. In case of bounded intervals $[\alpha_t, \beta_t]$, in order for the sets $\mathfrak{M}_t$ to be nonempty we assume that (compare with (6))

$$\mu_t \in [\alpha_t, \beta_t] \text{ and } \sigma_t^2 \leq (\beta_t - \mu_t)(\mu_t - \alpha_t),\ t = 1, ..., T. \tag{21}$$

The associated minimax problem supposes that although the set of associated marginal distributions may be "worst-case", the joint distribution will always be a product measure (i.e. the demand will be *independent across stages*). We conclude that the static formulation for the distributionally robust multistage news vendor problem may be formulated as follows:

$$\min_{\pi \in \Pi} \max_{Q \in \mathfrak{M}} \mathbb{E}_Q \left\{ \sum_{t=1}^{T} \rho^{t-1} \left[ c_t\big(x_t(y_t) - y_t\big) + \Psi_t\big(x_t(y_t), D_t\big) \right] \right\}, \tag{22}$$

where $\Pi$ is the set of admissible policies defined previously in Section 3.1. Of course, if the set $\mathfrak{M} = \{Q\}$ is a singleton, then formulation (22) coincides with formulation (15) taken with respect to the distribution $Q = Q_1 \times \cdots \times Q_T$ of the demand vector $(D_1, ..., D_T)$.

We note that very little is known about the set of optimal policies for problem (22), as this problem does not enjoy a dynamic-programming formulation along the lines of (17). Note also that as in Section 3.1, we only consider policies for which the amount ordered in stage $t$ is a function of $y_t$, and will discuss this choice in greater detail in the following section.

**3.2.2. Dynamic formulation** We now consider the so-called *dynamic* formulation, in which the policy maker recomputes her policy choices after each stage. Let us think on what it means for the policy maker to "recompute" her optimal policy at the start of the final stage $T$. As she cannot change her past decisions, the only policy decision she still has to make is the determination of the function $x_T$. However, she now has knowledge of $y_T$, the realized inventory level at the start of stage $T$, which she can incorporate into her minimax computations. We note that here we are faced with the modeling question of how to reconcile the use of $y_T$'s realized value in performing one's minimax computations (i.e. selecting the worst-case demand in stage $T$), which depends on past realizations of demand, with the fact that in the static formulation, the policy maker supposed that the demand in each stage was independent.

We take a conditional approach in line with the conditioning framework used in previous formulations of time consistent risk averse problems (cf., Ruszczyński and Shapiro 2006). Similar to (17) we write the following dynamic programming equations

$$V_t(y) = \min_{x \geq y} \left\{ c_t(x-y) + \max_{Q_t \in \mathfrak{M}_t} \mathbb{E}_{Q_t}[\Psi_t(x, D_t) + \rho V_{t+1}(x - D_t)] \right\}, \tag{23}$$

$t = 1, ..., T$, with $V_{T+1}(\cdot) \equiv 0$.

In analogy with (18), the dynamic programming equations (23) naturally define a set of optimal policies through the relation

$$x_t(y) \in \mathfrak{Y}_t(y) := \arg\min_{x \geq y} \left\{ c_t(x-y) + \max_{Q_t \in \mathfrak{M}_t} \mathbb{E}_{Q_t}[\Psi_t(x, D_t) + \rho V_{t+1}(x - D_t)] \right\}, \ t = 1, ..., T, \tag{24}$$

with associated optimal value given by $V_1(y_1)$. We define the set of optimal policies for the dynamic formulation to be those policies $\pi = (x_1, \ldots, x_T) \in \Pi$ such that $x_t(y) \in \mathfrak{Y}_t(y)$ for all $y \in \mathbb{R}, t = 1, \ldots, T$. We now briefly comment on the meaning of the dynamic programming equations (24), and also refer the reader to (Shapiro 2011, section 3.4) for a more detailed discussion.

Note that for any given policy $\pi \in \Pi$, the objective function $\sum_{t=1}^T \rho^{t-1} \big[ c_t\big(x_t(y_t) - y_t\big) + \Psi_t\big(x_t(y_t), D_t\big) \big]$ for the static formulation is a deterministic function $Z = Z(D_1, ..., D_T)$ of the demand process. The value associated with this policy under the static formulation is $\max_{Q \in \mathfrak{M}} \mathbb{E}_Q[Z]$.

For the random variable $Z = Z(D_1, ..., D_T)$ and probability measure $Q$ of $(D_1, ..., D_T)$ we have

$$\mathbb{E}_Q[Z] = \mathbb{E}_{Q|D_0} \Big[ \mathbb{E}_{Q|D_{[1]}} \big[ \cdots \mathbb{E}_{Q|D_{[T-1]}}[Z] \big] \Big], \tag{25}$$

where for uniformity of the notation we assume that $D_0$ is deterministic. Thus

$$\max_{Q \in \mathfrak{M}} \mathbb{E}_Q[Z] \leq \max_{Q \in \mathfrak{M}} \mathbb{E}_{Q|D_0} \Big[ \max_{Q \in \mathfrak{M}} \mathbb{E}_{Q|D_{[1]}} \big[ \cdots \max_{Q \in \mathfrak{M}} \mathbb{E}_{Q|D_{[T-1]}}[Z] \big] \Big]. \tag{26}$$

Since probability measures $Q \in \mathfrak{M}$ are of the form $Q = Q_1 \times \cdots \times Q_T$, we can write (26) as

$$\max_{Q \in \mathfrak{M}} \mathbb{E}_Q[Z] \leq \max_{Q_1 \in \mathfrak{M}_1} \mathbb{E}_{Q_1} \Big[ \max_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2} \big[ \cdots \max_{Q_T \in \mathfrak{M}_T} \mathbb{E}_{Q_T}[Z(D_1, ..., D_T)] \big] \Big]. \tag{27}$$

Then informally, the value of the policy $\pi$ under the static formulation coincides with the left-hand side of (27), while the value of the policy $\pi$ under the dynamic formulation coincides with the right-hand side of (27). Furthermore, the nested structure of the right-hand side of (27) allows the dynamic formulation to be solved by dynamic programming, with optimal policies defined by (24).

**Observation 4** *It follows from (27) that the optimal value $V_1(y_1)$ of the dynamic formulation is always greater than or equal to the optimal value of the static problem (22). Moreover, in principle the inequality can be strict. This can be explained as follows.*

*For the sake of simplicity let $T=2$. Let us fix some policy $\pi\in\Pi$, and let $Z_\pi = Z_\pi(D_1,D_2)$ denote the corresponding objective function. Clearly for any $Q_2\in\mathfrak{M}_2$ the following inequality holds*

$$\mathbb{E}_{Q_2}[Z_\pi(D_1,D_2)] \le \max_{Q_2'\in\mathfrak{M}_2} \mathbb{E}_{Q_2'}[Z_\pi(D_1,D_2)] \tag{28}$$

*for all $D_1$, and hence for any $Q_1\in\mathfrak{M}_1$ and $Q_2\in\mathfrak{M}_2$,*

$$\mathbb{E}_{Q_1\times Q_2}[Z_\pi(D_1,D_2)] \le \mathbb{E}_{Q_1}\left[\max_{Q_2'\in\mathfrak{M}_2} \mathbb{E}_{Q_2'}[Z_\pi(D_1,D_2)]\right]. \tag{29}$$

*By taking maximum of both sides of (29) with respect to $Q_1\in\mathfrak{M}_1$ and $Q_2\in\mathfrak{M}_2$ we obtain (27). Suppose that the maximum in the right hand side of (28) is attained at some distribution $\bar{Q}_2=\bar{Q}_2(D_1)$, which in general depends on $D_1$. If $\bar{Q}_2$ does not depend on $D_1$, then for $Q_2=\bar{Q}_2$ the equality holds in (28) and hence the corresponding inequality (27) holds as equality. If, on the other hand, $\bar{Q}_2(D_1)$ does depend on $D_1$, then for any $Q_2\in\mathfrak{M}_2$ the inequality (28) is strict for some values of $D_1$. Therefore the possible dependence of $\bar{Q}_2(D_1)$ on $D_1$ is a reason why the inequality (27) can be strict. We will elaborate on this further in the following sections.*

*In particular, suppose that $\mathfrak{M}_1=\{Q_1\}$ is a singleton with $Q_1=\frac{1}{2}\delta_0+\frac{1}{2}\delta_1$. Then the associated cost under the static formulation equals*

$$\max_{Q_2\in\mathfrak{M}_2} \mathbb{E}_{Q_1\times Q_2}[Z_\pi(D_1,D_2)] = \tfrac{1}{2}\max_{Q_2\in\mathfrak{M}_2}\left(\int Z_\pi(0,\tau)dQ_2(\tau) + \int Z_\pi(1,\tau)dQ_2(\tau)\right),$$

*while the associated cost under the dynamic formulation equals*

$$\mathbb{E}_{Q_1}\left[\max_{Q_2\in\mathfrak{M}_2} \mathbb{E}_{Q_2}[Z_\pi(D_1,D_2)]\right] = \tfrac{1}{2}\max_{Q_2\in\mathfrak{M}_2}\int Z_\pi(0,\tau)dQ_2(\tau) + \tfrac{1}{2}\max_{Q_2\in\mathfrak{M}_2}\int Z_\pi(1,\tau)dQ_2(\tau).$$

*Thus the value of the dynamic formulation is strictly greater than that of the static formulation if*

$$\begin{aligned}\min_{\pi\in\Pi}\left(\max_{Q_2\in\mathfrak{M}_2}\int Z_\pi(0,\tau)dQ_2(\tau) + \max_{Q_2\in\mathfrak{M}_2}\int Z_\pi(1,\tau)dQ_2(\tau)\right) > \\ \min_{\pi\in\Pi}\max_{Q_2\in\mathfrak{M}_2}\left(\int Z_\pi(0,\tau)dQ_2(\tau) + \int Z_\pi(1,\tau)dQ_2(\tau)\right).\end{aligned}$$

We note that it follows from the nature of the dynamic programming equations and our definitions that in the dynamic formulation, under any optimal policy, the amount ordered in stage $t$ is a function of $y_t$. As we wish to compare the set of optimal policies of the static and dynamic formulations, it makes sense to consider only those static policies which share this property, which (as described previously) is the approach we take here. However, we note that loosening this restriction, and considering a more general family of policies for the static problem, could potentially

introduce qualitatively different types of optimal policies for the static formulation, and this may be an interesting direction for future research.

Let us recall the following classical definition of a basestock policy.

**Definition 3.1** *A policy $\pi = (x_1(y_1), x_2(y_2), \ldots, x_T(y_T))$, for either the static or dynamic formulation, is said to be a* basestock *policy if there exist constants $x_t^* \in \mathbb{R}$, $t = 1, \ldots, T$, such that $x_t(y) = \max\{y, x_t^*\}$ for all $y \in \mathbb{R}$ and $t = 1, \ldots, T$.*

**Observation 5** *It is well known that the classical formulation of the multistage inventory problem always has an optimal basestock policy (e.g., Zipkin (2000)). Furthermore, it follows from the convexity of the cost-to-go functions $V_t$, defined in (23), that the distributionally robust dynamic formulation possesses an optimal basestock policy (cf., Ahmed, Cakmak and Shapiro 2007).*

We note that the question of whether or not there exists such an optimal basestock policy for the static formulation is considerably more challenging, and will be central to our discussion of time consistency in the next section.

## 4. Time consistency

As discussed in Section 3.2, there is no apriori guarantee that the static formulation is equivalent to the corresponding dynamic formulation in the distributionally robust setting. Disagreement between these two formulations is undesirable from a policy perspective, as it suggests that a policy which was optimal when performing one's minimax computations before seeing any realized demand may no longer be optimal if one reperforms these computations at a later time. This general problem goes under the heading of time (in)consistency. Although first addressed within the economics community, the issue of time (in)consistency has recently started to receive attention in the stochastic and robust optimization communities (e.g., Artzner et al. (2007), Carpentier et al. (2012), Huang et al. (2011), Riedel (2004), Ruszczyński (2010), Shapiro (2012) and references therein).

A well known quotation of Bellman (1957), coming from his pioneering work on dynamic programming, asserts that: *"An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision."* In a somewhat more precise form this principle has been formulated, e.g., in Carpentier et al. (2012) as: *"The decision maker formulates an optimization problem at time $t_0$ that yields a sequence of optimal decision rules for $t_0$ and for the following time steps $t_1, \ldots, t_N = T$. Then, at the next time step $t_1$, he formulates a new problem starting at $t_1$ that yields a new sequence of optimal decision rules from time steps $t_1$ to $T$. Suppose the process*

*continues until time $T$ is reached. The sequence of optimization problems is said to be dynamically consistent if the optimal strategies obtained when solving the original problem at time $t_0$ remain optimal for all subsequent problems."*

From a conceptual point of view this is quite natural - an optimal solution obtained by solving the problem at the first stage should remain optimal from the point of view of later stages. The setting in which one re-optimizes at each stage coincides precisely with the dynamic formulation for the distributionally robust multistage news vendor problem, while the "problem at time $t_0$" coincides naturally with the static formulation. Intuitively, this motivates us to define an optimal policy $\bar{\pi} = (\bar{x}_1, \ldots, \bar{x}_T)$ of the static problem to be time consistent if $\bar{\pi}$ is also optimal for the dynamic formulation. However, we will have to take some care here, to rule out the following situation. Note that by definition, $\bar{\pi}$ is also an optimal policy for the dynamic formulation iff $\bar{x}_t(y) \in \mathfrak{Y}_t(y)$ for all $y \in \mathbb{R}$. However, certain values of $y$ may be irrelevant under the static formulation, e.g. values of $y$ with zero probability, and a policy may take any action whatsoever for such values of $y$ while remaining optimal for the static formulation. Of course, we do not wish to declare a problem time inconsistent for such trivial reasons, and thus wish to compare policies only for "relevant" values of $y$. We now make this precise. For a policy $\pi \in \Pi$, let $\Xi^{\pi} := \arg\max_{Q \in \mathfrak{M}} \mathbb{E}_Q \left\{ \sum_{t=1}^{T} \rho^{t-1} \left[ c_t\big(x_t(y_t) - y_t\big) + \Psi_t\big(x_t(y_t), D_t\big) \right] \right\}$. If $\Xi^{\pi} = \emptyset$, we let $\hat{\Xi}^{\pi}$ denote the set of all sequences of distributions $\{Q_n, n \geq 1\}$ s.t. $Q_n \in \mathfrak{M}$ for all $n$, and $\lim_{n\to\infty} \mathbb{E}_{Q_n} \left\{ \sum_{t=1}^{T} \rho^{t-1} \left[ c_t\big(x_t(y_t) - y_t\big) + \Psi_t\big(x_t(y_t), D_t\big) \right] \right\}$ equals the value of $\pi$ under the static formulation. For $\pi \in \Pi$, $Q \in \Xi^{\pi}$, and $t \in 1, \ldots, T$, let $Y_t^{\pi,Q}$ denote the family of all measurable sets $A \in \mathbb{R}$ s.t. $Q(y_t \in A) = 1$, i.e. the probability that $y_t \in A$ under policy $\pi$, if $D_{[T]}$ is distributed as $Q$, equals one. For $\pi \in \Pi$ s.t. $\Xi^{\pi} = \emptyset$, and a sequence $\{Q_n, n \geq 1\} \in \hat{\Xi}^{\pi}$, let $Y_t^{\pi,\{Q_n\}}$ denote the family of all measurable sets $A \in \mathbb{R}$ s.t. $\lim_{n\to\infty} Q_n(y_t \in A) = 1$. We now combine the above ideas to formally define time consistency in our setting, and will use this definition for the remainder of the paper.

**Definition 4.1** *Suppose that $\bar{\pi} = (\bar{x}_1, \ldots, \bar{x}_T)$ is an optimal policy of the static problem. If $\Xi^{\bar{\pi}} \neq \emptyset$, we say that $\bar{\pi}$ is* time consistent *if there exists $Q \in \Xi^{\bar{\pi}}$ s.t. for all $t \in [1,T]$, there exists $A_t \in Y_t^{\bar{\pi},Q}$ s.t. $\bar{x}_t(y) \in \mathfrak{Y}_t(y)$ for all $y \in A_t$. If $\Xi^{\bar{\pi}} = \emptyset$, we say that $\bar{\pi}$ is* time consistent *if there exists $\{Q_n, n \geq 1\} \in \hat{\Xi}^{\bar{\pi}}$ s.t. for all $t \in [1,T]$, there exists $A_t \in Y_t^{\bar{\pi},\{Q_n\}}$ s.t. $\bar{x}_t(y) \in \mathfrak{Y}_t(y)$ for all $y \in A_t$. We say that the static problem is* weakly time consistent *if it possesses at least one time consistent optimal policy. We say that the static problem is* strongly time consistent *if its every optimal policy is time consistent.*

That is, weak time consistency amounts to the condition that the intersection of the sets of optimal policies of the static and dynamic formulations is non-empty. Strong time consistency is

equivalent to requiring that the entire set of optimal policies for the static formulation is contained in the set of optimal policies of the dynamic formulation.

If the set $\mathfrak{M}$ is a singleton, and hence we are back to the classical formulation, then it is well-known that strong time consistency follows (we already mentioned this in Section 3.1). Interestingly, it follows from (Shapiro 2011, section 4.2.2) that if one only has information about the support $\mathcal{I}_t$ and first moment $\mu_t \in \mathcal{I}_t$ of demand at each stage, and $\mathcal{I}_t = [\alpha_t, \beta_t]$ is bounded for all $t$, then the corresponding distributionally robust multistage inventory problem is strongly time consistent. In that case the dynamic formulation reduces to the static formulation (in an appropriate sense), as it can be shown to follow from convexity that in all stages, the "worst-case" demand distribution is independent of previous demand. As we will see, the question of time consistency becomes much more involved when one is also given second moment information.

### 4.1. Sufficient conditions for time consistency

In this section, we provide simple sufficient conditions for the time consistency of Problem (22). We begin by providing a different (but equivalent) formulation for Problem (22), in which all relevant instances of $y_t$ are rewritten in terms of the appropriate $x_t$ functions, as this will clarify the precise structure of the relevant cost-to-go functions. As a notational convenience, let $c_{T+1} = 0$, in which case we define

$$\hat{\Psi}_t(x_t, d_t) := (c_t - \rho c_{t+1})x_t + b_t[d_t - x_t]_+ + h_t[x_t - d_t]_+, \; t = 1, ..., T. \tag{30}$$

Let us define the problem

$$\min_{\pi \in \Pi} \max_{Q \in \mathfrak{M}} \mathbb{E}_Q \left[ \sum_{t=1}^{T} \rho^{t-1} \hat{\Psi}_t\big(x_t(y_t), D_t\big) \right] - c_1 y_1 + \sum_{t=1}^{T-1} \rho^t c_{t+1} \mu_t. \tag{31}$$

Then it follows from a straightforward substitution and calculation that

**Observation 6** *Problem* (22) *and Problem* (31) *are equivalent, i.e. each policy* $\pi \in \Pi$ *has the same value under both formulations.*

We now derive a lower bound for any policy, which intuitively comes from allowing the policy maker to reselect her inventory at the start of each stage, at no cost. As it turns out, this bound is "realizable" when the set of basestock levels is monotone increasing. For $x \in \mathbb{R}$, let us define

$$\eta_t(x) := \max_{Q_t \in \mathfrak{M}_t} \mathbb{E}_{Q_t}[\hat{\Psi}_t(x, D_t)], \;\; \Gamma_t^x := \arg\max_{Q_t \in \mathfrak{M}_t} \mathbb{E}_{Q_t}[\hat{\Psi}_t(x, D_t)], \tag{32}$$

and let

$$\begin{aligned} \hat{\eta}_t &:= \min_{x \in \mathbb{R}} \eta_t(x) = \min_{x \in \mathbb{R}} \max_{Q_t \in \mathfrak{M}_t} \mathbb{E}_{Q_t}[\hat{\Psi}_t(x, D_t)], \\ \hat{\Gamma}_t &:= \arg\min_{x \in \mathbb{R}} \eta_t(x) = \arg\min_{x \in \mathbb{R}} \max_{Q_t \in \mathfrak{M}_t} \mathbb{E}_{Q_t}[\hat{\Psi}_t(x, D_t)]. \end{aligned} \tag{33}$$

For $j \geq 1$, and probability measures $Q_1, \ldots, Q_j$, denote $\otimes_{t=1}^{j} Q_t := Q_1 \times \ldots \times Q_j$. Then we have the following.

**Lemma 4.1** *Suppose that the sets $\Gamma_t^x, \hat{\Gamma}_t$ are non-empty for all $x \in \mathbb{R}$, $t = 1,...,T$. Let us fix any $\pi = (x_1,\ldots,x_T) \in \Pi$, and $i \geq 0$. Then for any given $Q_1 \in \mathfrak{M}_1,\ldots,Q_i \in \mathfrak{M}_i$, there exist $Q_{i+1} \in \mathfrak{M}_{i+1},\ldots,Q_T \in \mathfrak{M}_T$ such that*

$$\mathbb{E}_{\otimes_{t=1}^T Q_t}\big[\hat{\Psi}_t\big(x_t(y_t), D_t\big)\big] \geq \hat{\eta}_t \text{ for all } t \geq i+1. \tag{34}$$

*Furthermore, the optimal value of Problem* (22) *is at least* $\sum_{t=1}^T \rho^{t-1}\hat{\eta}_t - c_1 y_1 + \sum_{t=1}^{T-1} \rho^t c_{t+1}\mu_t$.

*Proof* Suppose $i \in \{0,...,T\}$ and $Q_1,\ldots,Q_i$ are fixed. We now prove that (34) holds for all $t \geq i+1$, and proceed by induction. Our particular induction hypothesis will be that there exist $Q_{i+1},\ldots,Q_{i+n}$ such that

$$\mathbb{E}_{\otimes_{t=1}^{i+n} Q_t}\big[\hat{\Psi}_t\big(x_t(y_t), D_t\big)\big] \geq \hat{\eta}_t \text{ for all } t \in [i+1, i+n]. \tag{35}$$

We first treat the base case $n = 1$. It follows from Jensen's inequality, and the independence structure of the measures in $\mathfrak{M}$, that for any $Q_{i+1} \in \mathfrak{M}_{i+1}$,

$$\mathbb{E}_{\otimes_{t=1}^{i+1} Q_t}\big[\hat{\Psi}_{i+1}\big(x_{i+1}(y_{i+1}), D_{i+1}\big)\big] \geq \mathbb{E}_{Q_{i+1}}\big[\hat{\Psi}_{i+1}\big(\mathbb{E}_{\otimes_{t=1}^{i} Q_t}[x_{i+1}(y_{i+1})], D_{i+1}\big)\big].$$

Taking $Q_{i+1}$ to be any element of $\Gamma_{i+1}^{\mathbb{E}_{\otimes_{t=1}^{i} Q_t}[x_{i+1}(y_{i+1})]}$ ($\Gamma_1^{x_1(y_1)}$ if $i = 0$) completes the proof for $n = 1$.

Now, suppose the induction holds for some $n$. It again follows from Jensen's inequality, and the independence structure of the measures in $\mathfrak{M}$, that for any $Q_{i+n+1} \in \mathfrak{M}_{i+n+1}$,

$$\mathbb{E}_{\otimes_{t=1}^{i+n+1} Q_t}\big[\hat{\Psi}_{i+n+1}\big(x_{i+n+1}(y_{i+n+1}), D_{i+n+1}\big)\big] \geq \mathbb{E}_{Q_{i+n+1}}\big[\hat{\Psi}_{i+n+1}\big(\mathbb{E}_{\otimes_{t=1}^{i+n} Q_t}[x_{i+n+1}(y_{i+n+1})], D_{i+n+1}\big)\big].$$

Taking $Q_{i+n+1}$ to be any element of $\Gamma_{i+n+1}^{\mathbb{E}_{\otimes_{t=1}^{i+n} Q_t}[x_{i+n+1}(y_{i+n+1})]}$ completes the induction, and the proof, where the second part of the lemma follows by letting $i = 0$. $\square$

We now show that the bound of Lemma 4.1 is "realizable" when the set of basestock levels is monotone increasing, and that in this case the associated basestock policy is optimal for both the static and dynamic formulations. In particular, in this setting, the associated basestock policy is time consistent, and thus the static problem is weakly time consistent.

**Proposition 4.1** *Suppose there exists $x^* = (x_1^*,...,x_T^*)$ such that $y_1 \leq x_1^*$, $\{x_t^*,\ t = 1,\ldots,T\}$ is nondecreasing, and $x_t^* \in \hat{\Gamma}_t$ for all $t = 1,...,T$. Also suppose $\mathcal{I}_t \subset \mathbb{R}_+$ for all $t = 1,...,T$. Then the basestock policy $\pi$ for which $x_t(y) = \max\{y, x_t^*\}$ for all $y \in \mathbb{R}$, is an optimal policy for both the static and dynamic formulations, and attains value $\sum_{t=1}^T \rho^{t-1}\hat{\eta}_t - c_1 y_1 + \sum_{t=1}^{T-1} \rho^t c_{t+1}\mu_t$. Consequently, this basestock policy is time consistent, and the static problem is weakly time consistent.*

*Proof* Note that under these assumptions, if policy $\pi$ is implemented under the dynamic formulation, then w.p.1 $x_t(y_t) = x_t^*$ for all $t = 1, ..., T$. It then follows from a straightforward induction that $\pi$ is an optimal policy for the dynamic formulation, and w.p.1, for all $t = 2, ..., T$,

$$V_t(y_t) = \hat{\eta}_t - c_t x_{t-1}^* + c_t D_{t-1} + \sum_{s=t+1}^{T} \rho^{s-t}(\hat{\eta}_s + c_s \mu_{s-1}),$$

and

$$V_1(y_1) = \sum_{t=1}^{T} \rho^{t-1}\hat{\eta}_t - c_1 y_1 + \sum_{t=1}^{T-1} \rho^t c_{t+1}\mu_t.$$

Combining with Lemma 4.1 and Observation 5 completes the proof. $\square$

We now show that under additional assumptions, which intuitively correspond to requiring that there is a unique optimal basestock policy, and in this policy all basestock constants equal zero, the static problem becomes strongly time consistent.

**Theorem 3** *Suppose that* $b_t' := b_t - c_t + \rho c_{t+1} > 0$, $h_t' := h_t + c_t - \rho c_{t+1} > 0$, $\sigma_t, \mu_t > 0$, $\mathcal{I}_t = \mathbb{R}_+, t = 1, \ldots, T$, $y_1 = 0$, *and*

$$\frac{\sigma_t^2}{\mu_t^2} > \frac{b_t'}{h_t'},\ t = 1, ..., T. \tag{36}$$

*Then the set of optimal policies for the static problem is exactly the set of policies*

$$\Pi^0 := \big\{\pi = (x_1, \ldots, x_T) \in \Pi : x_1(y_1) = 0, x_t(z) = 0\, for\, all\; z \le 0\; and\; t \in [1, T]\big\},$$

*and the static problem is strongly time consistent.*

The result of the above theorem shows (under the specified conditions for the involved parameters) that if $y_1 = 0$ and variance at each stage is sufficiently large, then the basestock policy which always orders up to exactly zero is the optimal policy for both the static and dynamic formulations (compare with condition (i) of Theorem 1). This can be interpreted as asserting that if the variance at each stage is sufficiently high, then ordering up to any strictly positive amount only gives "nature" (which is picking the worst-case distribution) more "freedom" to select a distribution which yields higher costs. Although we do not formally prove that the conditions of Theorem 3 are tight for strong time consistency, our later examples show that if one deviates slightly from the conditions of Theorem 3, it becomes possible for the static problem to lose this property.

*Proof of Theorem 3* Let $\Pi^{opt}$ denote the set of optimal policies for the static problem. It follows from Theorem 1(i) and Proposition 4.1 that $\Pi^0 \subseteq \Pi^{opt}$, and every policy $\pi \in \Pi^0$ is time consistent. Thus to prove the theorem, it suffices to demonstrate that $\Pi^0 = \Pi^{opt}$, and we begin by showing that $\bar{\pi} = (\bar{x}_1, \ldots, \bar{x}_T) \in \Pi^{opt}$ implies $\bar{x}_1(y_1) = 0$. Indeed, it follows from Lemma 4.1 that $\bar{\pi} \in \Pi^{opt}$ implies

$$\max_{Q \in \mathfrak{M}_1} \mathbb{E}_Q\big[\hat{\Psi}_1\big(\bar{x}_1(y_1), D_1\big)\big] = \hat{\eta}_1 = b_1\mu_1.$$

That $\bar{x}_1(y_1)$ must equal 0 then follows from Theorem 1.

We now show that $\bar{\pi} \in \Pi^{opt}$ implies $\bar{x}_2(z) = 0$ for all $z \le 0$. Suppose for contradiction that there exists $z' \le 0$ s.t. $\bar{x}_2(z') \ne 0$. It is easily verified that there exists $Q_1 \in \mathfrak{M}_1$ such that $Q_1(-z') > 0$, and consequently for this choice of $Q_1$, $\bar{x}_2(y_2)$ is not a.s. equal to 0. We conclude from Proposition 2.1 that there exists $Q_2 \in \mathfrak{M}_2$ such that

$$\mathbb{E}_{Q_1 \times Q_2}\left[\hat{\Psi}_2\big(\bar{x}_2(y_2), D_2\big)\right] > \hat{\eta}_2 = b_2\mu_2.$$

As we have already demonstrated that $\bar{x}_1(y_1) = 0$, and $Q_1 \in \mathfrak{M}_1$, we conclude that

$$\mathbb{E}_{Q_1}\left[\hat{\Psi}_1\big(\bar{x}_1(y_1, D_1)\big)\right] = \hat{\eta}_1 = b_1\mu_1.$$

Combining with Lemma 4.1 then yields a contradiction. The proof that $\bar{x}_t(z) = 0$ for all $z \le 0$ and $t \ge 3$ follows from a nearly identical argument, and we omit the details. $\square$

### 4.2. Further study of time (in)consistency

In this section, we show that the static problem is neither strongly nor weakly time consistent in general. Furthermore, our examples demonstrate that the question of weak and strong time consistency can be quite subtle in this setting. Throughout this section, we will let $\Pi_s^{opt}$ denote the set of all optimal policies for the corresponding static problem, and $\Pi_d^{opt}$ denote the set of all optimal policies for the corresponding dynamic problem.

**4.2.1. Example when the static problem is not weakly time consistent** In this section, we explicitly provide an example for which the static problem is not weakly time consistent, showing that in general, the static and dynamic formulations need not have a common optimal policy. Furthermore, for this example, the static and dynamic formulations have different optimal values.

Let us define $y_1 = 10$, $\rho = 1$,

$$\mathcal{I}_1 = [1,3], \quad \mu_1 = 2, \quad \sigma_1 = 1, \quad c_1 = 0, \quad b_1 = 2, \quad h_1 = 2,$$

$$\mathcal{I}_2 = \mathbb{R}_+, \quad \mu_2 = 8, \quad \sigma_2 = 2, \quad c_2 = 0, \quad b_2 = 1, \quad h_2 = 1.$$

Let $\tilde{\Pi}_s$ denote the set of policies $\tilde{\pi} = (\tilde{x}_1, \tilde{x}_2)$ such that $\tilde{x}_1(10) = 10$, $\tilde{x}_2(9) = 9$, $\tilde{x}_2(7) = 7$, and $\tilde{\Pi}_d$ denote the set of policies $\tilde{\pi} = (\tilde{x}_1, \tilde{x}_2)$ such that $\tilde{x}_1(10) = 10$, $\tilde{x}_2(9) = 9$, $\tilde{x}_2(7) = 8$. Note that the set $\tilde{\Pi}_s$ specifies values of $x_2(y_2)$ only for $y_2 = 9$ and $y_2 = 7$. As we will see in Lemma 4.2, other values of $y_2$ are irrelevant for the static formulation regarding optimality.

**Theorem 4** $\Pi_s^{opt} = \tilde{\Pi}_s$, *and the optimal value of the static problem is 18. On the other hand,* $\Pi_d^{opt} \subseteq \tilde{\Pi}_d$, *and the optimal value of the dynamic problem is* $17 + \frac{\sqrt{5}}{2} > 18$. *Consequently, the static problem is not weakly time consistent, and the static and dynamic problems have different optimal values.*

We first characterize the set of optimal policies for the static problem.

**Lemma 4.2** $\Pi_s^{opt} = \tilde{\Pi}_s$*, and the static problem has optimal value 18.*

*Proof* It follows from Observation 1 that $\mathfrak{M}_1$ consists of the single probability measure $Q_1$ such that $Q_1(1) = Q_1(3) = \frac{1}{2}$. Let $D_1$ denote a r.v. distributed as $Q_1$. Note that for any policy $\pi = (x_1, x_2) \in \Pi$, one has that $x_1(y_1) = x_1(10) \geq 10$. Consequently, $\mathsf{Pr}(x_1(y_1) \geq D_1) = 1$, and $|x_1(y_1) - D_1| = x_1(y_1) - D_1$ w.p.1. It then follows from a straightforward calculation that the cost of any policy $\pi = (x_1, x_2) \in \Pi$ under the static formulation equals

$$2x_1(10) - 4 + \max_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\Big[\tfrac{1}{2}\Big(\big|x_2\big(x_1(10) - 1\big) - D_2\big| + \big|x_2\big(x_1(10) - 3\big) - D_2\big|\Big)\Big]. \tag{37}$$

Let $\bar{\pi} = (\bar{x}_1, \bar{x}_2)$ denote any optimal policy for the static problem, i.e. $\bar{\pi} \in \Pi_s^{opt}$. Then it follows from (37) and a straightforward contradiction argument that

$$\bar{x}_1(10) = 10. \tag{38}$$

Combining (37) and (38), we conclude that

$$\big(\bar{x}_2(9), \bar{x}_2(7)\big) \in \underset{(x,y): x \geq 9, y \geq 7}{\arg\min} \max_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\big[\tfrac{1}{2}\big(|x - D_2| + |y - D_2|\big)\big]. \tag{39}$$

Furthermore, it follows from Lemma 4.1 and Theorem 1 that

$$\min_{(x,y): x \geq 9, y \geq 7} \max_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\big[\tfrac{1}{2}\big(|x - D_2| + |y - D_2|\big)\big] \quad \geq \quad \max_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\big[|8 - D_2|\big] \quad = \quad 2. \tag{40}$$

Noting that

$$\tfrac{1}{2}\big(|9 - D_2| + |7 - D_2|\big) = 1 + \max(-D_2 + 7, 0, D_2 - 9),$$

it then follows from a straightforward calculation and Theorem 2 that

$$\max_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\big[\tfrac{1}{2}\big(|9 - D_2| + |7 - D_2|\big)\big] = 2. \tag{41}$$

Combining the above, we conclude that $\tilde{\Pi}_s \subseteq \Pi_s^{opt}$. Also, it then follows from a straightforward calculation that the static problem has optimal value 18.

We now prove that $\tilde{\Pi}_s = \Pi_s^{opt}$. Indeed, suppose for contradiction that there exists some optimal policy $\hat{\pi} = (\hat{x}_1, \hat{x}_2) \notin \tilde{\Pi}_s$. In that case, it follows from (38) and (39) that $\frac{1}{2}\big(\hat{x}_2(9) + \hat{x}_2(7)\big) > 8$. However, it then follows from Jensen's inequality, Theorem 1, and (40) that

$$\max_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\big[\tfrac{1}{2}\big(|\hat{x}_2(9) - D_2| + |\hat{x}_2(7) - D_2|\big)\big] \quad \geq \quad \max_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\big[\big|\tfrac{1}{2}\big(\hat{x}_2(9) + \hat{x}_2(7)\big) - D_2\big|\big] \quad > \quad 2.$$

Combining with (40) and (41) yields a contradiction, completing the proof $\square$

We now characterize the set of optimal policies for the dynamic problem.

**Lemma 4.3** $\Pi_d^{opt} \subseteq \tilde{\Pi}_d$, *and the dynamic problem has optimal value* $17+\frac{\sqrt{5}}{2}$.

*Proof* Let $\bar{\pi}=(\bar{x}_1,\bar{x}_2)$ denote any optimal policy for the dynamic problem, i.e. $\bar{\pi}\in\Pi_d^{opt}$. Then it again follows from a straightforward contradiction argument that

$$\bar{x}_1(10)=10. \tag{42}$$

It then follows from (24) that

$$\bar{x}_2(9)\in \arg\min_{x\geq 9}\max_{Q_2\in\mathfrak{M}_2}\mathbb{E}_{Q_2}[|x-D_2|],$$

and

$$\bar{x}_2(7)\in \arg\min_{x\geq 7}\max_{Q_2\in\mathfrak{M}_2}\mathbb{E}_{Q_2}[|x-D_2|].$$

The lemma then follows from Theorem 1 and a straightforward calculation $\square$

Combining Lemmas 4.2 and 4.3 completes the proof of Theorem 4.

**4.2.2. Example when the static problem is weakly time consistent, but *not* strongly time consistent** In this section, we explicitly provide an example showing that it is possible for the static problem to be weakly time consistent, but not strongly time consistent. In particular, the static and dynamic formulations have a common optimal basestock policy $\pi^*$, with associated basestock constants $x_1^*, x_2^*$, satisfying the conditions of Proposition 4.1, yet the static problem has other non-trivial optimal policies which are suboptimal for the dynamic formulation. The intuitive explanation is as follows. In the static formulation, one can leverage the randomness in the realization of $D_1$ to construct a policy $\pi'$ such that with positive probability $x_2'(y_2)$ is slightly below $x_2^*$, and with the remaining probability is slightly above $x_2^*$. Since in the static formulation "nature" cannot observe the realized inventory in stage 2 before selecting a worst-case distribution, it turns out that such a policy incurs the same cost as $\pi'$ under the static formulation. Alternatively, this policy is suboptimal in the dynamic formulation, as the adversary can first see exactly how the inventory level deviated from $\pi^*$, and exploit this to achieve a strictly higher cost. We note that in this example, even though the static problem is not strongly time consistent, both formulations have the same optimal value, as dictated by Proposition 4.1.

Let us define $y_1=0$, $\rho=1$,

$$\mathcal{I}_1=[1,3],\quad \mu_1=2,\quad \sigma_1=1,\quad c_1=0,\quad b_1=1,\quad h_1=1,$$

$$\mathcal{I}_2=\mathbb{R}_+,\quad \mu_2=10,\quad \sigma_2=1,\quad c_2=0,\quad b_2=1,\quad h_2=1.$$

Then we prove the following.

**Theorem 5** *The static problem is weakly time consistent, but not strongly time consistent.*

We first prove that the static problem is weakly time consistent.

**Lemma 4.4** *The static problem is weakly time consistent, and both the static and dynamic problems have optimal value 2.*

*Proof* Note that

$$\hat{\Psi}_1(x_1, d_1) = |x_1 - d_1|, \quad \hat{\Psi}_2(x_2, d_2) = |x_2 - d_2|.$$

It follows from Observation 1 that $\mathfrak{M}_1$ consists of the single probability measure $Q_1$ such that $Q_1(1) = Q_1(3) = \frac{1}{2}$. It follows from Theorem 1 and a straightforward calculation that

$$\hat{\Gamma}_1 = [1,3] \quad , \quad \hat{\Gamma}_2 = 10 \quad , \quad \hat{\eta}_2 = 1.$$

Combining the above with Proposition 4.1, we conclude that the basestock policy $\pi$ such that $x_1(y) = \max\{3, y\}$, and $x_2(y) = \max\{10, y\}$ for all $y \in \mathbb{R}$, is optimal for both the static and dynamic problems, which have common optimal value 2. $\square$

We now prove that the static problem is not strongly time consistent. In particular, consider the policy $\pi' = (x'_1, x'_2)$ such that

$$x'_1(y) = \max\{3, y\}, \quad \text{and} \quad x'_2(y) = \begin{cases} 9.9, & \text{if } y \le 0, \\ \max\{10.1, y\}, & \text{otherwise.} \end{cases} \tag{43}$$

**Lemma 4.5** *The policy $\pi' \in \Pi_s^{opt}$, but $\pi' \notin \Pi_d^{opt}$. Consequently, the static problem is not strongly time consistent.*

*Proof* We first show that $\pi' \in \Pi_s^{opt}$. It follows from a straightforward calculation that the cost of $\pi'$ under the static formulation equals

$$\mathbb{E}_{Q_1}|3 - D_1| + 0.1 + \max_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2} \max\left\{9.9 - D_1, 0, D_1 - 10.1\right\}. \tag{44}$$

It is easily verified that the conditions of Theorem 2 are met, and we may apply Theorem 2 to conclude that $\arg\max_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2} \max\left\{9.9 - D_1, 0, D_1 - 10.1\right\}$ is the probability measure $Q_2$ such that $Q_2(9) = \frac{1}{2}$, $Q_2(11) = \frac{1}{2}$. It follows that the value of expression in (44) equals 2, and we conclude that $\pi' \in \Pi_s^{opt}$, completing the proof.

We now show that $\pi' \notin \Pi_d^{opt}$. Suppose, for contradiction, that $\pi' \in \Pi_d^{opt}$. It then follows from a straightforward calculation that

$$9.9 \in \arg\min_{x \ge 0} \max_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}[|x - D_2|] \tag{45}$$

However, it follows from Theorem 1 that the right-hand side of (45) is the singleton $\{10\}$, completing the proof $\square$

Combining Lemmas 4.4 and 4.5 completes the proof of Theorem 5.

**4.2.3. Example when the static problem is strongly time consistent, but the two formulations have a *different* optimal value** In this section, we explicitly provide an example showing that it is possible for the static problem to be strongly time consistent, yet for the two formulations to have different optimal values. We note that, although it is expected that there will be settings where the two formulations have different optimal values, it is somewhat surprising that this is possible even when the two formulations have the same set of optimal policies.

Let us define $y_1 = 0$, $\rho = 1$,

$$\mathcal{I}_1 = [1,3], \quad \mu_1 = 2, \quad \sigma_1 = 1, \quad c_1 = 0, \quad b_1 = 0, \quad h_1 = 0,$$

$$\mathcal{I}_2 = \mathbb{R}_+, \quad \mu_2 = 100, \quad \sigma_2 = 5, \quad c_2 = 2, \quad b_2 = 1, \quad h_2 = 1.$$

Let $\tilde{\Pi}$ denote the set of policies $\tilde{\pi} = (\tilde{x}_1, \tilde{x}_2)$ such that $\tilde{x}_1(0) = 102$, $\tilde{x}_2(101) = 101$, $\tilde{x}_2(99) = 99$. Then we prove the following.

**Theorem 6** $\Pi_s^{opt} = \tilde{\Pi}$, *and the static problem is strongly time consistent. However, the optimal value of the static problem equals 5, while the optimal value of the dynamic problem equals* $\sqrt{26} > 5$.

We first characterize the set of optimal policies for the static problem.

**Lemma 4.6** $\Pi_s^{opt} = \tilde{\Pi}$, *and the static problem has optimal value 5.*

*Proof* It follows from Observation 1 that $\mathfrak{M}_1$ consists of the single probability measure $Q_1$ such that $Q_1(1) = Q_1(3) = \frac{1}{2}$. In this case, the cost of any policy $\pi = (x_1, x_2) \in \Pi$ under the static formulation equals

$$\max_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\Big[\mathbb{E}_{Q_1}\Big[2\Big(x_2\big(x_1(0) - D_1\big) - \big(x_1(0) - D_1\big)\Big) + \big|x_2\big(x_1(0) - D_1\big) - D_2\big|\Big]\Big]. \tag{46}$$

We now prove that for any policy $\bar{\pi} = (\bar{x}_1, \bar{x}_2) \in \Pi_s^{opt}$, one has that

$$\bar{x}_2\big(\bar{x}_1(0) - 1\big) = \bar{x}_1(0) - 1 \quad \text{and} \quad \bar{x}_2\big(\bar{x}_1(0) - 3\big) = \bar{x}_1(0) - 3. \tag{47}$$

Indeed, note that w.p.1, it follows from the triangle inequality that

$$\begin{aligned}
&2\Big(x_2\big(x_1(0) - D_1\big) - \big(x_1(0) - D_1\big)\Big) + \big|x_2\big(x_1(0) - D_1\big) - D_2\big| \\
&= 2\Big(x_2\big(x_1(0) - D_1\big) - \big(x_1(0) - D_1\big)\Big) + \big|x_2\big(x_1(0) - D_1\big) - \big(x_1(0) - D_1\big) + \big(x_1(0) - D_1\big) - D_2\big| \\
&\geq 2\Big(x_2\big(x_1(0) - D_1\big) - \big(x_1(0) - D_1\big)\Big) + \big|\big(x_1(0) - D_1\big) - D_2\big| - \big|x_2\big(x_1(0) - D_1\big) - \big(x_1(0) - D_1\big)\big| \\
&= x_2\big(x_1(0) - D_1\big) - \big(x_1(0) - D_1\big) + \big|x_1(0) - D_1 - D_2\big|.
\end{aligned} \tag{48}$$

Now, suppose for contradiction that (47) does not hold. It follows that

$$\mathbb{E}_{Q_1}\big[x_2\big(x_1(0) - D_1\big) - \big(x_1(0) - D_1\big)\big] > 0,$$

and combining with (48), we conclude that (46) is strictly greater than

$$\max_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\left[\mathbb{E}_{Q_1}\left[\left|x_1(0) - D_1 - D_2\right|\right]\right]. \tag{49}$$

Noting that (49) is the cost incurred by some policy satisfying (47) completes the proof.

We now complete the proof of the lemma. It suffices from the above to prove that

$$\arg\min_{x_1 \in \mathbb{R}_+} \max_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\left[\tfrac{1}{2}\big(|x_1 - 1 - D_2| + |x_1 - 3 - D_2|\big)\right] = \{102\}. \tag{50}$$

It follows from a straightforward calculation that as long as $x_1 \geq 3$, $(x_1 - 100)(104 - x_1) \leq 25$ and $x_1 - 2 - \big((x_1 - 2 - 100)^2 + 25\big)^{\frac{1}{2}} \geq 0$, which holds for all $x_1 \in [100, 104]$, the conditions of Theorem 2 are met. We may thus apply Theorem 2 to conclude that for all $x_1 \in [100, 104]$,

$$\max_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\left[\tfrac{1}{2}\big(|x_1 - 1 - D_2| + |x_1 - 3 - D_2|\big)\right] \tag{51}$$

has the unique optimal solution $\hat{Q}_2$ such that

$$\hat{Q}_2\Big(x_1 - 2 - \big((x_1 - 2 - 100)^2 + 25\big)^{\frac{1}{2}}\Big) = 25\Big(25 + \big(x_1 - 2 - \big((x_1 - 2 - 100)^2 + 25\big)^{\frac{1}{2}} - 100\big)^2\Big)^{-1},$$

and

$$\hat{Q}_2\Big(x_1 - 2 + \big((x_1 - 2 - 100)^2 + 25\big)^{\frac{1}{2}}\Big) = 1 - 25\Big(25 + \big(x_1 - 2 - \big((x_1 - 2 - 100)^2 + 25\big)^{\frac{1}{2}} - 100\big)^2\Big)^{-1}.$$

It then follows from a straightforward calculation that for $x_1 \in [100, 104]$, (51) has the value

$$g(x_1) := \big(x_1^2 - 204x_1 + 10429\big)^{\frac{1}{2}}.$$

It is easily verified that $g$ is a strictly convex function on $[100, 104]$, $g$ has its unique minimum on that interval at the point 102, and $g(102) = 5$. The desired result then follows from the fact that (51) is a convex function of $x_1$ on $\mathbb{R}$. $\square$

We now prove that the static problem is strongly time consistent.

**Lemma 4.7** *The static problem is strongly time consistent, and the optimal value of the dynamic problem equals* $\sqrt{26}$.

*Proof* First, we note that as in the static setting, any policy $\bar{\pi} = (\bar{x}_1, \bar{x}_2) \in \Pi_d^{opt}$ also satisfies (47). The proof is very similar to that used for the static case, and we omit the details. To prove the lemma, it thus suffices to prove that

$$\arg\min_{x_1 \in \mathbb{R}_+} \left(\tfrac{1}{2} \max_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\left[|x_1 - 1 - D_2|\right] + \tfrac{1}{2} \max_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\left[|x_1 - 3 - D_2|\right]\right) = \{102\}. \tag{52}$$

It is easily verified that for all $x_1 \in [100, 104]$, we may apply Theorem 1 to conclude that

$$\max_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\left[|x_1 - 1 - D_2|\right] = \left((x_1 - 101)^2 + 25\right)^{\frac{1}{2}},$$
$$\max_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\left[|x_1 - 3 - D_2|\right] = \left((x_1 - 103)^2 + 25\right)^{\frac{1}{2}}.$$

We conclude that for all $x_1 \in [100, 104]$,

$$\tfrac{1}{2} \max_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\left[|x_1 - 1 - D_2|\right] + \tfrac{1}{2} \max_{Q_2 \in \mathfrak{M}_2} \mathbb{E}_{Q_2}\left[|x_1 - 3 - D_2|\right] \tag{53}$$

equals

$$g(x_1) := \tfrac{1}{2}\left(\left((x_1 - 101)^2 + 25\right)^{\frac{1}{2}} + \left((x_1 - 103)^2 + 25\right)^{\frac{1}{2}}\right). \tag{54}$$

It is easily verified that $g(x)$ is a strictly convex function of $x$ on $[100, 104]$, $g$ has its unique minimum on that interval at the point 102, and $g(102) = \sqrt{26}$. The desired result then follows from the fact that (53) is a convex function of $x_1$ on $\mathbb{R}$. $\square$

Combining Lemmas 4.6 and 4.7 completes the proof of Theorem 6.

## 5. Conclusion

In this paper, we considered a minimax approach to managing an inventory under distributional uncertainty. In particular, we studied the associated multistage distributionally robust optimization problem, when only the mean, variance and distribution support are known for the demand at each stage. Our contributions were four-fold. First, we gave a novel definition for time consistency in this context. More precisely, we defined two formulations for the relevant optimization problem. In the static formulation, the policy-maker cannot recompute her policy after observing realized demand. In the dynamic formulation, she is allowed to reperform her minimax computations at each stage. If these two formulations have a common optimal policy, we defined the static problem to be weakly time consistent. If all optimal policies of the static problem are also optimal for the dynamic problem, we defined the static problem to be strongly time consistent.

Next, we gave sufficient conditions for weak and strong time consistency. Intuitively, our sufficient condition for weak time consistency coincided with the existence of an optimal basestock policy in which the basestock constants are monotone increasing. Our sufficient condition for strong time consistency could be interpreted in two ways. On the one hand, strong time consistency holds if the unique optimal basestock policy for the dynamic formulation is to order-up to 0 at each stage. Alternatively, we saw that this condition also has an interpretation in terms of requiring that the demand variances are sufficiently large relative to their respective appropriate means.

Third, we gave a series of examples of two-stage problems exhibiting interesting and counterintuitive time (in)consistency properties, showing that the question of time consistency can be quite

subtle in this setting. In particular, we showed that: (i) the static problem could fail to be weakly time consistent, (ii) the static problem could be weakly but not strongly time consistent, and (iii) the static problem could be strongly time consistent even when the two formulations had different optimal values. Interestingly, this stands in contrast to the analogous setting in which only the mean and support of the demand distribution is known at each stage, for which it is known that such time inconsistency cannot occur Shapiro (2012).

Finally, as it was necessary for our investigations of time consistency, we extended Scarf's well-known solution to the single-stage distributionally robust newsvendor problem for convex, continuous piecewise affine function with exactly two pieces to a class of convex, continuous, piecewise affine functions with three pieces. We further note that this technique should generalize to any number of pieces, and such results may be of independent interest.

Our work leaves many interesting directions for future research. The general question of time consistency remains poorly understood. Furthermore, our work has shown that this question can be quite subtle. For the particular model we consider here, it would be interesting to develop a better understanding of the tightness of our sufficient conditions. Our examples w.r.t. time (in)consistency demonstrated that surprising phenomena can occur here, and a more thorough investigation of exactly which types of phenomena can and cannot happen, and when, would be beneficial. It is also an intriguing question to understand how much the two formulations can differ in optimal value and policy, even when time inconsistency occurs, along the lines of Huang et al. (2011), Agrawal et al. (2012). On a related note, it is largely open to develop a broader understanding of the optimal solution to the static problem, or even approximately optimal solutions, as well as related algorithms. We note that one potential avenue for understanding this question is to consider a third formulation, in which the adversary can select any demand distribution $(D_1, \ldots, D_T)$ whatsoever, so long as the marginal distribution $Q_t \in \mathfrak{M}_t$ for all $t$, as this formulation always yields an optimal value even larger than that of the dynamic formulation. Of course, it is also an open challenge to understand the question of time consistency more broadly, and more generally to understand the relationship between different ways to model optimization under uncertainty.

## References

Agrawal, S., Y. Ding, A. Saberi, Y. Ye. 2012. Price of correlations in stochastic optimization. *Operations Research.* **60**(1) 150-162.

Ahmed, S., U. Cakmak, A. Shapiro. 2007. Coherent risk measures in inventory problems. *European Journal of Operational Research.* **182** 226–238.

Artzner, P., F. Delbaen, J.M. Eber, D. Heath, H. Ku. 2007. Coherent multiperiod risk adjusted values and Bellmans principle. *Annals of Operations Research.* **152** 5–22.

Bellman, R.E. (1957). *Dynamic Programming.* Princeton University Press, Princeton, NJ.

Bonnans, J. F., A. Shapiro. 2000. *Perturbation Analysis of Optimization Problems.* Springer-Verlag, New York.

Carpentier, P., J.P. Chancelier, G. Cohen, M. De Lara and P. Girardeau. 2012. Dynamic consistency for stochastic optimal control problems, *Annals of Operations Research*, Online First.

Choi, S., A. Ruszczyński, Y. Zhao. 2011. A multiproduct risk-averse newswendor with law-invariant coherent measures of risk. *Operations Research*, **59**, 346–354.

Dupacová, J. 1987. The minimax approach to stochastic programming and an illustrative application. *Stochastics.* **20** 73–88.

Dupacová, J. 2001. Stochastic programming: Minimax approach. In: Ch. A. Floudas and P.M. Pardalos (Eds.), *Encyclopedia of Optimization.* **5** 327–330. Kluwer Academic Publishers.

Edgeworth, F. 1888. The mathematical theory of banking. *Royal Statistical Society*, 51, 113–127.

Gallego, G., I. Moon. 1993. The distribution free newsboy problem: review and extensions. *Journal of the Operational Research Society.* **44**(8) 825–834.

Gallego, G., K. Katircioglu, B. Ramachandran. 2007. Inventory management under highly uncertain demand. *Operations Research Letters.* **35**(3) 281-289.

Huang, P., D. Iancu, M. Petrik, D. Subramanian. 2011. The Price of Dynamic Inconsistency for Distortion Risk Measures. arXiv preprint arXiv:1106.6102.

Isii, K., 1963. On sharpness of Tchebycheff-type inequalities, *Ann. Inst. Stat. Math.,* 14, 185–197.

Moments in mathematics, H.J. Landau (ed.), Proc. Sympos. Appl. Math., 37, Amer. Math. Soc., Providence, RI, 1987.

Prékopa, A. 1995. *Stochastic Programming.* Kluwer Academic Publishers.

Riedel, F. 2004. Dynamic coherent risk measures. *Stochastic Processes and their Applications.* **112** 185–200.

Ruszczyński, A., A. Shapiro. 2006. Conditional risk mappings, *Mathematics of Operations Research*, **31** 544–561.

Ruszczyński, A. 2010. Risk-averse dynamic programming for Markov decision processes. *Mathematical Programming, Series B.* **125** 235–261.

Scarf, H. 1958. A min-max solution of an inventory problem. K. Arrow, ed *Studies in the Mathematical Theory of Inventory and Production.* Stanford University Press, Stanford, CA, 201–209.

Shapiro, A. 2001. On duality theory of conic linear problems. In: Miguel A. Goberna and Marco A. Lopez (Eds.), *Semi-Infinite Programming: Recent Advances.* Kluwer Academic Publishers 135–165.

Shapiro, A., D. Dentcheva and A. Ruszczyński. 2009. *Lectures on Stochastic Programming: Modeling and Theory.* SIAM, Philadelphia.

Shapiro, A. 2009. On a time consistency concept in risk averse multistage stochastic programming. *Operations Research Letters*. **37** 143–147.

Shapiro, A. 2011. *Topics in Stochastic Programming*, CORE Lecture Series, Universite Catholique de Louvain.

Shapiro, A. 2012. Minimax and risk averse multistage stochastic programming. *European Journal of Operational Research*. **219** 719–726.

Žáčková, J. 1966. On minimax solution of stochastic linear programming problems. *Časopis Pěstování Matematiky*. **91** 423–430.

Zipkin, P.H. 2000. *Foundations of Inventory Management*. McGraw-Hill, Boston.

# 6. Appendix

## 6.1. Proof of Theorem 1

*Proof of Theorem 1* We first compute the value of $\psi(x)$ for all $x \in \mathbb{R}$, and proceed by a case analysis. First, suppose $x < 0$. In this case, $\mathbb{E}_Q[\Psi(x,D)] = cx + b(\mu - x)$ for all $Q \in \mathfrak{M}$, and thus

$$\psi(x) = cx + b(\mu - x). \tag{55}$$

Now, suppose $x \geq 0$. Then it is easily verified that

$$\psi(x) = cx + \frac{(h-b)(x-\mu)}{2} + \frac{b+h}{2} \max_{Q \in \mathfrak{M}} \mathbb{E}_Q\left[|x - D|\right]. \tag{56}$$

Hence to compute $\psi(x)$, it suffices to solve $\max_{Q \in \mathfrak{M}} \mathbb{E}_Q\left[|x-D|\right]$, and we proceed by a case analysis. Recall that $\mathfrak{f}(z) := \left((z-\mu)^2 + \sigma^2\right)^{\frac{1}{2}}$ for all $z \in \mathbb{R}$.

First, suppose $x \geq \frac{\mu^2+\sigma^2}{2\mu}$. Let us define $\bar{\lambda} = (\bar{\lambda}_0, \bar{\lambda}_1, \bar{\lambda}_2)$ s.t.

$$\bar{\lambda}_0 := \frac{1}{2}\left(x^2 \mathfrak{f}^{-1}(x) + \mathfrak{f}(x)\right), \quad \bar{\lambda}_1 := -x\mathfrak{f}^{-1}(x), \quad \bar{\lambda}_2 := \frac{1}{2}\mathfrak{f}^{-1}(x),$$

and let $\bar{g}(d) := \bar{\lambda}_0 + \bar{\lambda}_1 d + \bar{\lambda}_2 d^2$ for all $d \in \mathbb{R}$. Then it follows from a straightforward calculation that $\bar{g}(d)$ and $|x-d|$ are tangent at $\bar{d}_1 := x - \mathfrak{f}(x)$ and $\bar{d}_2 := x + \mathfrak{f}(x)$, and consequently $\bar{g}(d) \geq |x-d|$ for all $d \in \mathbb{R}_+$. Hence $\bar{\lambda}$ is feasible for the dual Problem (10). Also, as $x \geq \frac{\mu^2+\sigma^2}{2\mu}$ implies $\bar{d}_1 \geq 0$, it is easily verified that the probability measure $\bar{Q}$ such that

$$\bar{Q}(\bar{d}_1) = \sigma^2\left(\sigma^2 + \left(x - \mathfrak{f}(x) - \mu\right)^2\right)^{-1}, \quad \bar{Q}(\bar{d}_2) = 1 - \sigma^2\left(\sigma^2 + \left(x - \mathfrak{f}(x) - \mu\right)^2\right)^{-1}$$

is feasible for the primal Problem (7). It follows from Proposition 2.2 that $\bar{Q}$ is an optimal primal solution. Combining the above and simplifying the relevant algebra, we conclude that in this case

$$\psi(x) = \psi_1(x) := c\mu + \frac{b+h}{2}\mathfrak{f}(x) - \frac{b-h-2c}{2}(x-\mu). \tag{57}$$

Alternatively, suppose $x \in [0, \frac{\mu^2+\sigma^2}{2\mu})$. Let us define $\hat{\lambda} = (\hat{\lambda}_0, \hat{\lambda}_1, \hat{\lambda}_2)$ s.t.

$$\hat{\lambda}_0 := x \quad , \quad \hat{\lambda}_1 := 1 - 4x\mu(\mu^2+\sigma^2)^{-1} \quad , \quad \hat{\lambda}_2 := 2x\big(\mu(\mu^2+\sigma^2)^{-1}\big)^2,$$

and let $\hat{g}(d) := \hat{\lambda}_0 + \hat{\lambda}_1 d + \hat{\lambda}_2 d^2$ for all $d \in \mathbb{R}$. Then it follows from a straightforward calculation that $\hat{g}(d)$ and $|x-d|$ are tangent at $\hat{d}_1 := \mu^{-1}(\mu^2+\sigma^2)$, and intersect at $\hat{d}_2 := 0$, with $\hat{g}'(0) \geq -1$. It follows that $\hat{g}(d) \geq |x-d|$ for all $d \in \mathbb{R}_+$. Hence $\hat{\lambda}$ is feasible for the dual Problem (10). Also, it is easily verified that the probability measure $\hat{Q}$ such that

$$\hat{Q}(\hat{d}_1) = \mu^2(\mu^2+\sigma^2)^{-1}, \quad \hat{Q}(\hat{d}_2) = 1 - \mu^2(\mu^2+\sigma^2)^{-1}$$

is feasible for the primal Problem (7). It follows from Proposition 2.2 that $\hat{Q}$ is an optimal primal solution. Combining the above and simplifying the relevant algebra, we conclude that in this case

$$\psi(x) = \psi_2(x) := \frac{(h+c)\sigma^2 - (b-c)\mu^2}{\mu^2+\sigma^2}x + b\mu. \tag{58}$$

We now use the above to complete the proof of the theorem. Note that since by assumption $b > c$, it follows from (55) that $\arg\min_{x\in\mathbb{R}} \psi(x) \subseteq \mathbb{R}_+$. Recall that $\kappa = \frac{b-h-2c}{b+h}$. Furthermore, our assumptions, i.e. $b > c, h+c > 0$, imply that $|\kappa| < 1$. Let $\chi := \mu + \kappa\sigma(1-\kappa^2)^{-\frac{1}{2}}$. It follows from a straightforward calculation that $\psi_1$ is a strictly convex function on $\mathbb{R}$, and $\psi_1(\chi) = 0$, i.e. $\psi_1$ is strictly decreasing on $(-\infty, \chi)$, and strictly increasing on $(\chi, \infty)$. Furthermore, it follows from a similar calculation that

$$\frac{\sigma^2}{\mu^2} - \frac{b-c}{h+c} \quad \text{is the same sign as} \quad \frac{\mu^2+\sigma^2}{2\mu} - \chi. \tag{59}$$

We now proceed by a case analysis. First, suppose $\frac{\sigma^2}{\mu^2} > \frac{b-c}{h+c}$. In this case, $\psi_2$ is a linear function with strictly positive slope, and thus $\arg\min_{x\in[0,\frac{\mu^2+\sigma^2}{2\mu}]} \psi(x) = \{0\}$. Furthermore, it follows from (59) that $\chi < \frac{\mu^2+\sigma^2}{2\mu}$, which implies that $\psi_1$ is strictly increasing on $[\frac{\mu^2+\sigma^2}{2\mu}, \infty)$. It follows from the continuity of $\psi$ that $\arg\min_{x\geq\frac{\mu^2+\sigma^2}{2\mu}} \psi(x) = \{\frac{\mu^2+\sigma^2}{2\mu}\}$. Combining the above, we conclude that $\arg\min_{x\in\mathbb{R}} \psi(x) = \{0\}$.

Next, suppose $\frac{\sigma^2}{\mu^2} < \frac{b-c}{h+c}$. In this case, $\psi_2$ is a linear function with strictly negative slope, and thus $\arg\min_{x\in[0,\frac{\mu^2+\sigma^2}{2\mu}]} \psi(x) = \{\frac{\mu^2+\sigma^2}{2\mu}\}$. Furthermore, it follows from (59) that $\chi > \frac{\mu^2+\sigma^2}{2\mu}$, which implies that $\arg\min_{x\geq\frac{\mu^2+\sigma^2}{2\mu}} \psi(x) = \{\chi\}$. Combining the above, we conclude that $\arg\min_{x\in\mathbb{R}} \psi(x) = \{\chi\}$.

Finally, suppose that $\frac{\sigma^2}{\mu^2} = \frac{b-c}{h+c}$. In this case, $\psi_2$ is a constant function, and thus $\arg\min_{x\in[0,\frac{\mu^2+\sigma^2}{2\mu}]} \psi(x) = [0, \frac{\mu^2+\sigma^2}{2\mu}]$. Furthermore, it follows from (59) that $\chi = \frac{\mu^2+\sigma^2}{2\mu}$, which implies

that $\arg\min_{x \geq \frac{\mu^2+\sigma^2}{2\mu}} \psi(x) = \{\frac{\mu^2+\sigma^2}{2\mu}\}$. Combining the above, we conclude that $\arg\min_{x\in\mathbb{R}} \psi(x) = [0, \frac{\mu^2+\sigma^2}{2\mu}]$.

Combining all of the above with another straightforward calculation completes the proof of the theorem. $\square$

### 6.2. Proof of Proposition 2.1

*Proof of Proposition 2.1* We know that $\delta_0$ is an optimal solution to Problem (9) with value $b\mu$. We now prove that it is the unique optimal solution. Let $\delta := \frac{\sigma^2}{\mu^2+\sigma^2}, \tau := \frac{\mu^2+\sigma^2}{\mu}$. Let $Q_2^*$ be the probability measure such that

$$Q_2^*(0) = \delta, \quad Q_2^*(\tau) = 1 - \delta.$$

Recall that $b - c > 0$, and $(h+c)\sigma^2 > (b-c)\mu^2$, which we denote by assumption A1. Note that the value of any feasible solution $Q_1$ to Problem (9) is at least $\mathbb{E}_{Q_1 \times Q_2^*}\Big[\Psi(D_1, D_2)\Big]$, which itself equals the sum of $c\mu$ and

$$\mathbb{E}_{Q_1}\bigg[\bigg(\delta\big((b-c)[0-D_1]_+ + (h+c)[D_1-0]_+\big) + (1-\delta)\big((b-c)[\tau-D_1]_+ + (h+c)[D_1-\tau]_+\big)\bigg)I(D_1>0)\bigg] \tag{60}$$

$$+\ \mathbb{E}_{Q_1}\bigg[\bigg(\delta\big((b-c)[0-D_1]_+ + (h+c)[D_1-0]_+\big) + (1-\delta)\big((b-c)[\tau-D_1]_+ + (h+c)[D_1-\tau]_+\big)\bigg)I(D_1<0)\bigg] \tag{61}$$

$$+\ \mathbb{E}_{Q_1}\bigg[\bigg(\delta\big((b-c)[0-D_1]_+ + (h+c)[D_1-0]_+\big) + (1-\delta)\big((b-c)[\tau-D_1]_+ + (h+c)[D_1-\tau]_+\big)\bigg)I(D_1=0)\bigg] \tag{62}$$

Note that if $P(D_1>0)>0$, then (60) is at least

$$\begin{aligned}
&\mathbb{E}\Big[\frac{\sigma^2}{\mu^2+\sigma^2}(h+c)D_1 + \frac{\mu^2}{\mu^2+\sigma^2}(b-c)(\frac{\mu^2+\sigma^2}{\mu} - D_1)\big|D_1>0\Big]P(D_1>0) \\
> \quad &\mathbb{E}\Big[\frac{\mu^2}{\mu^2+\sigma^2}(b-c)D_1 + \frac{\mu^2}{\mu^2+\sigma^2}(b-c)(\frac{\mu^2+\sigma^2}{\mu} - D_1)\big|D_1>0\Big]P(D_1>0) \quad \text{by A1} \\
= \quad &(b-c)\mu P(D_1>0).
\end{aligned} \tag{63}$$

Similarly, if $P(D_1<0)>0$, then (61) is at least

$$\begin{aligned}
&\mathbb{E}\Big[-\frac{\sigma^2}{\mu^2+\sigma^2}(b-c)D_1 + \frac{\mu^2}{\mu^2+\sigma^2}(b-c)(\frac{\mu^2+\sigma^2}{\mu} - D_1)\big|D_1<0\Big]P(D_1<0) \\
= \quad &\mathbb{E}\Big[(b-c)(\mu-D_1)\big|D_1<0\Big]P(D_1<0) \quad > \quad (b-c)\mu P(D_1<0).
\end{aligned} \tag{64}$$

Similarly, if $P(D_1=0)>0$, then (62) equals $(b-c)\mu P(D_1=0)$. Combining with (63), (64), and the fact that $\delta_0$ attains value $b\mu$ completes the proof. $\square$

### 6.3. Proof of Theorem 2

*Proof of Theorem 2* Recall that $\eta := \frac{1}{2}(c_1 + c_2)$, and $\mathfrak{f}(z) := \big((z-\mu)^2 + \sigma^2\big)^{\frac{1}{2}}$ for all $z \in \mathbb{R}$. Also, letting $h_1(d) := -d + c_1, h_2(d) := d - c_2$ for all $d \in \mathbb{R}$, we have that $\Psi(x,d) = \max\{h_1(d), 0, h_2(d)\}$ for all $d \in \mathbb{R}$. Let $Q$ be the probability measure described in (12), and $\lambda = (\lambda_0, \lambda_1, \lambda_2)$ the vector described in (13). Let $g(d) := \lambda_0 + \lambda_1 d + \lambda_2 d^2$. We now prove that $g(d) \geq \Psi(x,d)$ for all $d \in \mathbb{R}$. It follows from a straightforward calculation that $g(d)$ is tangent to $h_1(d)$ at $d_1 := \eta - \mathfrak{f}(\eta)$, and $g(d)$ is tangent to $h_2(d)$ at $d_2 := \eta + \mathfrak{f}(\eta)$. Thus $g(d) \geq \max\big(h_1(d), h_2(d)\big)$ for all $d \in \mathbb{R}$, and to prove the desired claim it suffices to demonstrate that $g(d) \geq 0$ for all $d \geq 0$. It is easily verified that for all $d \in \mathbb{R}$,

$$g(d) = \tfrac{1}{2}\mathfrak{f}^{-1}(\eta)(d-\eta)^2 + \tfrac{1}{2}\big(\mathfrak{f}(\eta) + c_1 - c_2\big). \tag{65}$$

Recall that

$$\frac{1}{4}(2\mu - 3c_1 + c_2)(3c_2 - c_1 - 2\mu) \leq \sigma^2,$$

which we denote by assumption A2. It follows from another straightforward calculation that assumption A2 is equivalent to requiring that $\frac{1}{2}\big(\mathfrak{f}(\eta) + c_1 - c_2\big) \geq 0$. Combining with (65), we conclude that A2 implies $g(d) \geq 0$ for all $d \in \mathbb{R}$, completing the proof that $g(d) \geq \Psi(x,d)$ for all $d \in \mathbb{R}$. Hence $\lambda$ is feasible for the dual Problem (10). Also, it is easily verified that $Q$ is feasible for the primal Problem (7). It follows from Proposition 2.2 that $Q$ is an optimal primal solution, and $\lambda$ is an optimal dual solution. That these optimal solutions are unique then follows from the second part of Proposition 2.2 and a straightforward contradiction argument. Combining the above and simplifying the relevant algebra completes the proof. □